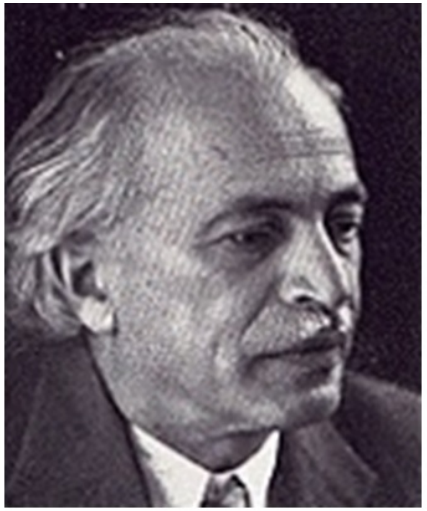
Boolean Functions: Influence, threshold and noise

Gil Kalai
Einstein Institute of Mathematics
Hebrew University of Jerusalem

Based on recent joint works with Jean Bourgain, Jeff Kahn, Guy Kindler, Nathan Keller, and Elchanan Mossel, and on older works with Itai Benjamini, Jean Bourgain, Ehud Friedgut, Jeff Kahn, Nati Linial, and Oded Schramm.

7ECM, Berlin, July 2016

Edmund Landau



Michael Fekete

Binyamin Amira

Abraham Fraenkel

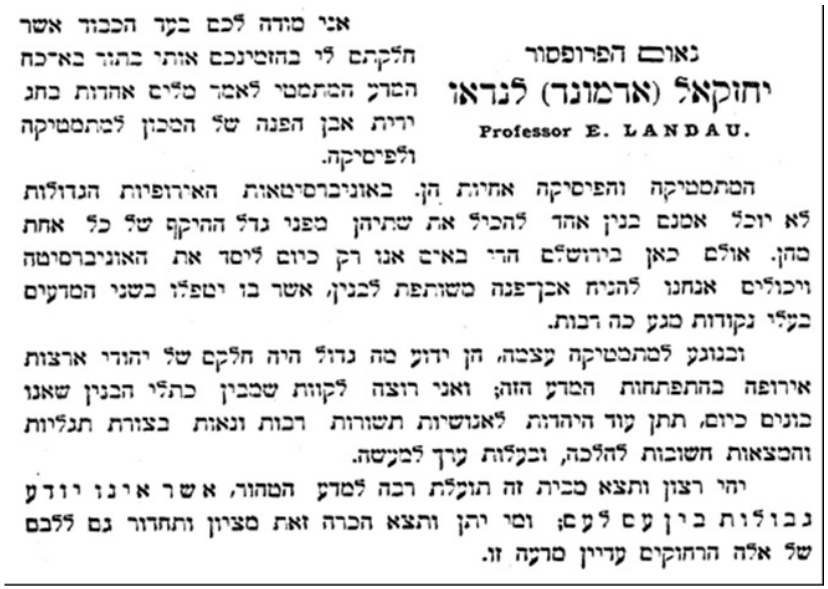
Haim Hanani

Landau's 1925 speech

Zionist Internationalism through Number Theory: Edmund Landau at the Opening of the Hebrew University in 1925

Leo Corry and Norbert Schappacher

Tel Aviv University and Université de Strasbourg



אני סומה לכם בעד הכבוד אשר
חלקתם לי בהזמינכם אותי בתוך באיכח
המדע המתמטי לאמר טלים אהדות בתג
ירית אבן הפנה של המכון למתמטיקה
ולפיסיקה.

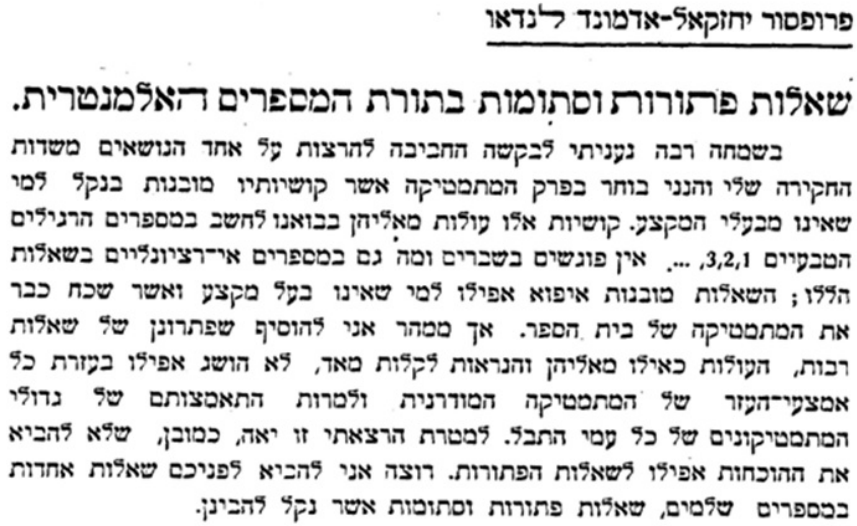
נאום הפרופסור
יחזקאל (אדמונד) לנדאו
Professor E. LANDAU.

המתמטיקה והפיסיקה אחיות הן. באוניברסיטאות האירופיות הגדולות
לא יוכל אנוש בנין אחד להכיל את שתיהן מפני גדל החרקף של כל אחת
מהן. אולם כאן בירושלים הרי באים אנו רק כיום ליסד את האוניברסיטה
ויכולים אנחנו להניח אבן־פנה משותפת לבנין, אשר בו יטפלו בשני המדעים
בעלי נקודות מגע כה רבות.
ובנוגע למתמטיקה עצמה, הן ידוע מה גדול היה חלקם של יהודי ארצות
אירופה בהתפתחות המדע הזה; ואני רוצה לקוות שבבין כתלי הבנין שאנו
בונים כיום, תתן עוד היהדות לאנושיות תשורות רבות ונאות בצורת תגליות
והמצאות חשובות להלכה, ובגלות ערך למעשה.
יהי רצון ותצא מכאן זה תועלת רבה למדע הטהור, אשר אינו יודע
גבולות בין עם לעם; ומי יתן ותצא הכרה זאת מציון ותתרור גם ללבם
של אלה הרחוקים עדיין מרעה זו.

... Mathematics and physics are sister disciplines... here in Jerusalem we are just founding a university today, and we can lay a common cornerstone for a building where the two sciences which have so many points of contact will be cultivated.

... May great benefit emerge from this house dedicated to pure science, **which does not know borders between people and people**. And may this awareness emerge from Zion and penetrate the hearts of all those who are still far from this view.

Landau's 1925 article



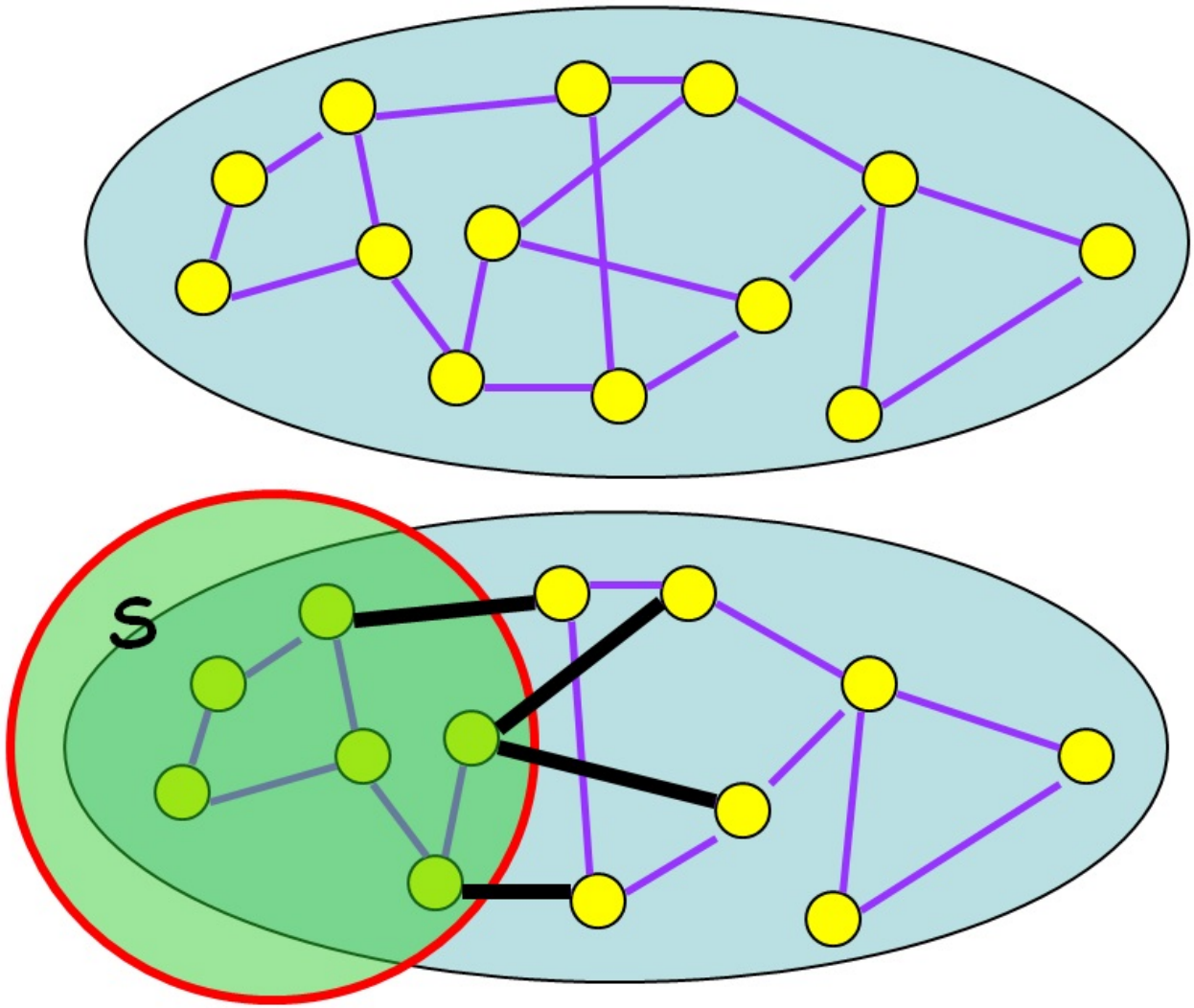
פרופסור יחזקאל-אדמונד לנדאו

שאלות פתורות וסתומות בתורת המספרים דאלמנטרית.

בשמחה רבה נעניתי לבקשה החביבה להרצות על אחד הנושאים משדות החקירה שלי והנני בוחר בפרק המתמטיקה אשר קושיותיו מובנות בנקל למי שאינו מבעלי הטקצע. קושיות אלו עולות מאליהן בבואנו לחשב במספרים הרגילים הטבעיים 3,2,1, ... אין פונשים בשברים ומה גם במספרים אי־רציונליים בשאלות הללו; השאלות מובנות אפוא אפילו למי שאינו בעל מקצע ואשר שכח כבר את המתמטיקה של בית הספר. אך טמהר אני להוסיף שפתרונן של שאלות רבות, העולות כאילו מאליהן והנראות לקלות מאד, לא הושג אפילו בעזרת כל אמצעי־העזר של המתמטיקה המודרנית ולטרות התאמצותם של גדולי המתמטיקונים של כל עמי התבל. למטרת הרצאתי זו יאה, כמובן, שלא להביא את ההוכחות אפילו לשאלות הפתורות. רוצה אני להביא לפניכם שאלות אחדות בטספרים שלמים, שאלות פתורות וסתומות אשר נקל להבינן.

At this number of twenty-three problems I want to stop, because twenty-three is a prime number, i.e., a very handsome number for us. I am certain that I should not fear to be asked by you, for what purpose does one deal with the theory of numbers and what application it may have. For we deal with science for the sake of it, and dealing with it was a solace in the days of internal and external war that as Jews and as Germans we fought and still fight today.

Introduction : Graphs and expansion



Graphs and expansion:

- Combinatorics: no small cuts, high connectivity
- Geometry: High isoperimetry
- Probability: Rapid convergence of random walk
- Spectral (Algebra/Analysis): Spectral gap for the Laplacian
- Computation: Many applications and connections

Zooming in: the discrete n-dimensional cube

The **discrete n-dimensional cube** Ω_n is the set of 0-1 vectors of length n.

Two vertices of the discrete cube are adjacent if they differ in one coordinate.

The uniform probability distribution on Ω_n is denoted by μ.

Expansion for the discrete cube

Theorem: The number of edges $e(A, \bar{A})$ between a set A of vertices of Ω_n and its complement $\bar{A}$ is at least $\min(|A|, |\bar{A}|)$.

Proof: (Giving a lower bound of $2^{-(n-1)}(|A| \cdot |\bar{A}|)$.
Let $E(A, \bar{A})$ be the set of edges between A and $\bar{A}$. Given two vertices x, y of the discrete cube we will consider the *canonical path* between x and y where we flip coordinates of disagreement from left to right.
Observations: 1) Every edge z, u of Ω_n belongs to precisely 2^{n-1} canonical paths.
2) Every canonical path from $x \in A$ to $y \in \bar{A}$ contains an edge from $E(A, \bar{A})$. $\square$

Part I: Influence

The discrete n-dimensional cube and Boolean functions

The **discrete n-dimensional cube** Ω_n is the set of 0-1 vectors of length n. The uniform probability distribution on Ω_n is denoted by μ.

A **Boolean function** f is a map from Ω_n to $\{0,1\}$. For a Boolean function f, $\mathbb{E}(f) = \mu(\{x : f(x) = 1\})$. Boolean functions are characteristic funcions of subsets of the discrete cube.

A Boolean function f is **monotone** if f cannot decrease when you switch a coordinate from 0 to 1.

Influence of variables on Boolean functions

Let

$$\sigma_k(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n) = (x_1, \ldots, x_{k-1}, 1-x_k, x_{k+1}, \ldots, x_n).$$

The influence of the kth variable on a Boolean function f is defined by:

$$I_k(f) = \mu(x \in \Omega_n, f(x) \neq f(\sigma_k(x))).$$

The total influence is defined by

$$I(f) = \sum_{k=1}^{n} I_k(f).$$

The edge-isoperimetric theorem

The total influence is defined by

$$I(f) = \sum_{k=1}^{n} I_k(f).$$

If $f = \chi_A$ then $I(f) = e(A, \bar{A})/2^{n-1}$.
In a similar way, $I_k(f) = 2^{-n+1} e_k(A, \bar{A})$, where $e_k(A, \bar{A})$ is the number of edges in the "k-direction" between A and $\bar{A}$. We proved that $I(f) \geq 4\mathbb{E}(f)(1 - \mathbb{E}(f))$.
A stronger inequality is:

$$I(f) \geq 2\mathbb{E}(f) \log_2(1/\mathbb{E}(f)).$$

KKL's theorem

Theorem (Kahn, Kalai, Linial, 1988) There exists a variable k such that

$$I_k(f) \geq C\mathbb{E}(f)(1-\mathbb{E}(f))\log n/n.$$

This result was conjectured by Michael Ben-Or and Nati Linial in 1985.

A sharp version (due to Talagrand)

$$\sum_{k=1}^{n} I_k(f)/\log(e/I_k(f)) \geq C\mathbb{E}(f)(1-\mathbb{E}(f)).$$

The Bernoulli measure

Let $p, 0 < p < 1$, be a real number. The probability measure μ_p is the product probability distribution whose marginals are given by $\mu_p(x_k = 1) = p$. Let $f : \Omega_n \to \{0, 1\}$ be a Boolean function.

$$\mathbb{E}_p(f) = \sum_{x \in \Omega_n} \mu_p(x) f(x) = \mu_p\{x : \ f(x) = 1\}.$$

The notions of influence, total influence, the edge-expansion theorem, and KKL's theorem all extend to the biased ($p \neq 1/2$) case.

Edge-expansion inequalities

Edge-expansion inequality ($p = 1/2$):

$$I(f) \geq 2\mathbb{E}(f)\log(1/\mathbb{E}(f)).$$

Edge-expansion inequality (general p):

$$I^p(f) \geq 2\mathbb{E}_p(f)\log_p(\mathbb{E}_p(f)).$$

Sketch of proof:
Induction on the number of coordinates. Let $\alpha = \mathbb{E}(f|x_n = 0)$ and $\beta = \mathbb{E}(f|x_n = 1)$. Based on the induction hypothesis one needs to prove

$$(1-p)\alpha\log_p\alpha + p\beta\log_p\beta + p(\beta-\alpha)$$

$$-((1-p)\alpha + p\beta)log_p((1-p)\alpha + p\beta) \geq 0.$$

This holds with equality when $\beta = \alpha$ and (by an easy calculation) the derivative of the l.h.s. w.r.t. β is nonnegative for all β.

Wanted: inverse theorems

Edge-expansion inequality (general p):

$$I^p(f) \geq 2\mathbb{E}_p(f)\log_p(1/\mathbb{E}_p(f)).$$

Problem 1: Inverse theorem for edge expansion: Understand the structure of Boolean functions for which

$$I^p(f) \leq K\mathbb{E}_p(f)\log_p(1/\mathbb{E}_p(f)).$$

For the case where both p and $\mathbb{E}_p(f)$ are bounded away from zero and one, Friedgut (1997) proved that such functions are "Juntas". Major theorems for the case where p is small but $\mathbb{E}_p(f)$ is bounded away from zero and one, was were proved by Friedgut (1999), Bourgain (1999), and Hatami (2010). The case where $\mathbb{E}_p(f)$ is small is wide open and very important.

Part II: Fourier

The Fourier expansion

Every real function $f : \Omega_n \to \mathbb{R}$ can be expressed in terms of the Fourier-Walsh basis.

$$f = \sum \{\hat{f}(S) W_S : \ S \subset [n]\},$$

where

$$W_S(x_1, x_2, \ldots, x_n) = (-1)^{\sum \{x_i : i \in S\}}.$$

Parseval: For real functions f, g on Ω_n,

$$< f, g >= \sum \hat{f}(S)\hat{g}(S).$$

(Here $< f, g >= \sum_{x \in \Omega_n} \mu(x) f(x) g(x).$)

Influence: the Fourier description

Let $f : \Omega_n \to \{0,1\}$ be a Boolean function and let $f = \sum\{\hat{f}(S)W_S : \ S \subset [n]\}$, be its Fourier expansion. We can easily obtain from Parseval's formula that:

$$\mathbb{E}(f) = \| f \|_2^2 = \sum \hat{f}^2(S).$$

The individual influences are given by,

$$I_k(f) = 4 \sum_{S \subset [n], k \in S} \hat{f}^2(S).$$

And summing up we obtain that the total influence is expressed in terms of the Fourier-Walsh expansion as follows:

$$I(f) = 4 \sum \hat{f}^2(S)|S|.$$

Strong spectral expansion

Let $t = \mathbb{E}(f)$ We saw that $t = \| f \|_2^2 = \sum_S \hat{f}^2(S)$ and that $I(f) = 4 \sum_S \hat{f}^2(S)|S|$. (This gives us again that $I(f) \geq 4t(1-t)$.)

Let's consider a probability distribution on subsets S of $\{1, 2, \ldots, n\}$, where the probability of a set S is $t^{-1}\hat{f}^2(S)$. We know that $I(f) \geq 2t \log_2(1/t)$ so when t is small the expected value of $|S|$ is at least $2 \log(1/t)$.

Can we say something stronger?

Can we say that *most* of the "energy" of f comes from levels where $|S| = \Omega(\log(1/t))$?

The anomaly of majority

For the majority function f, $I(f) = C\sqrt{n}$ yet most of the Fourier coefficients $\hat{f}(S)$ lies where $|S|$ is bounded!

Remarks:

1. This example shows that strong spectral expansion does not follow from expansion.
2. Benjamini, Kalai and Schramm (1999) proved that monotone functions f where most of the ℓ_2 norm is concentrated on S where $|S|$ is bounded must have positive (bounded away from zero) correlation with a weighted majority function.
3. The anomaly of majority is related to the fact that robust classical information and computation are at all possible, in our noisy world.

The noise operator

Given a real function f on the discrete cube $f = \sum\{\hat{f}(S)W_S : \ S \subset [n]\}$, we write

$$T_\rho(f) = \sum\{\hat{f}(S)(1-\rho)^{|S|}W_S : \ S \subset [n]\}.$$

The Bonami-Gross-Beckner hypercontractive inequality:

$$\| T_{1/2}(f) \|_2 \ \leq \ \| f \|_{5/4}.$$

Hypercontructivity and strong spectral expansion:

Theorem: Let f be a Boolean function with $\| f \|_2^2 = t$ Then

$$\sum \{\hat{f}^2(S) : 0 < |S| < \frac{1}{10} \log(1/t)\} \leq t^{1.1}.$$

Proof:

0. Parseval gives $I(f) = 4 \sum \hat{f}^2(S)|S|$.
1. Bonami-Gross-Beckner hypercontractive inequality.

$$\| T_{1/2}(f) \|_2^2 = \sum \hat{f}^2(S)(1/2)^{|S|} \leq \| f \|_{5/4}^2.$$

2. For Boolean functions the qth power of the q norm is the measure of the support and does not depend on q. If the support is small this means that the q-norm is very different from the r-norm if $r \neq q$.

Hypercontractivity, and the proof of KKL's theorem

The spectral edge-expansion theorem asserts that if g is a Boolean function, and $\mathbb{E}(g) = t$, then most of the contribution to $\| g \|_2^2$ comes from Fourier-Walsh coefficients $\hat{g}^2(S)$ where $|S| \geq C \log(1/t)$.

Let $f_k(x) = f(x) - f(\sigma_k(x))$. We can think about f_k as a partial derivative of f in the kth direction. Thus $\| f_k \|_2^2 = 4 I_k(f)$. The spectral edge-expansion theorem tells us that if $I_k(f) < 1/\sqrt{n}$,then the contribution of Fourier coefficients (squared) of f_k below the $\frac{1}{10} \log n$ level is negligible. This easily implies that the contribution to $\| f \|_2^2$ of Fourier coefficients $\hat{f}^2(S)$ where $0 < |S| < 0.1 \log n$ is negligible, and, therefore, the total influence of f is at least $C \log n \, \mathbb{E}(f)(1 - \mathbb{E}(f))$.

The Entropy Influence conjecture

Let $f : \Omega_n \to \{0,1\}$ be a Boolean function and let $f = \sum\{\hat{f}(S)W_S : \ S \subset [n]\}$, be its Fourier expansion. Define

$$H(f) = \sum \hat{f}^2(S) \log(1/\hat{f}^2(S)).$$

Problem 2 (The entropy Influence Conjecture Friedgut & Kalai, 1996): Prove that for some absolute constant C, for every Boolean function f,

$$I(f) \geq C \cdot H(f).$$

First passage percolation

Consider an infinite planar grid where every edge is assigned a length: 1 with probability $1/2$ and 2 with probability $1/2$ (independently).

Question: What is the variance $V(n)$ of the distance from $(0,0)$ to $(n,0)$.

Kesten proved (1993) that $V(n) = O(n)$ and Benjamini, Kalai, and Schramm (2003) proved that $V(n) = O(n/\log n)$. The proof is very similar to the proof of KKL's theorem. Unlike the KKL theorem this bound is not sharp.

Problem 3: Prove that $V(n) = O(n^{1-c})$ for some $c > 0$.

Part III: Threshold

Influence of variables (the biased case)

Let

$$\sigma_k(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n) = (x_1, \ldots, x_{k-1}, 1-x_k, x_{k+1}, \ldots, x_n).$$

The influence of the kth variable on a Boolean function f is defined by:

$$I_k^p(f) = \mu_p(x \in \Omega_n, f(x) \neq f(\sigma_k(x))).$$

$$I^p(f) = \sum_{k=1}^{n} I_k^p(f).$$

Russo's lemma

Russo's lemma: For a monotone Boolean function f,

$$d\mathbb{E}_p(f)/dp = I^p(f).$$

Very useful in percolation theory and other areas.

The **threshold interval** for a monotone Boolean function f is those values of p so that $\mu_p(f)$ is bounded away from 0 and 1. (Say $0.01 \leq \mu_p(f) \leq 0.99$.)

A typical application of Russo's lemma: If for every value p in the threshold interval $I^p(f)$ is large, then the threshold interval itself is short. This is called a **sharp threshold** phenomenon.

Invariance under transitive group

Theorem (Friedgut and Kalai, 1996): If a monotone Boolean function f with n variables is invariant under a transitive group of permutation of the variables, then its threshold interval is of length $O(1/\log n)$.

Proof: The relation

$$I^p(f) \geq C\mu_p(f)(1-\mu_p(f))\log n.$$

It follows from KKL's theorem (extended to the biased case) since all individual influences are the same.

Total influence under symmetry of primitive groups

For a transitive group of permutations $\Gamma \subset S_n$, let $I(\Gamma)$ be the minimum influence for a Γ-invariant function Boolean function with n variables.

Theorem (Bourgain and Kalai 1997): If Γ is primitive then one of the following possibilities hold.

- $$I(\Gamma) = \theta(\sqrt{n}),$$
- $$(logn)^{(k+1)/k-o(1)} \leq I(\Gamma) \leq C(\log n)^{(k+1)/k},$$

  Here $k \geq 1$ is an integer.
- $I(\Gamma)$ behaves like $(\log n)\mu(n)$, where $\mu(n) \leq \log\log n$ is growing in an arbitrary way.

The expectation threshold conjecture

Consider a random graph G in $G(n,p)$ and the graph property: G contains a copy of a specific graph H. (Note: H depends on n; a motivating example: H is a Hamiltonian cycle.) Let q be the minimal value for which the expected number of copies of H' in G is at least $1/2$ for every subgraph H' of H. Let p be the value for which the probability that G contains a copy of H is $1/2$.

Problem 4 (Conjecture by Kahn and Kalai from 2006): Show that

$$p/q = O(\log n).$$

The conjecture can be vastly extended to general Boolean functions. Connectivity shows: the $\log n$ factor cannot be removed.

Two slides on Percolation

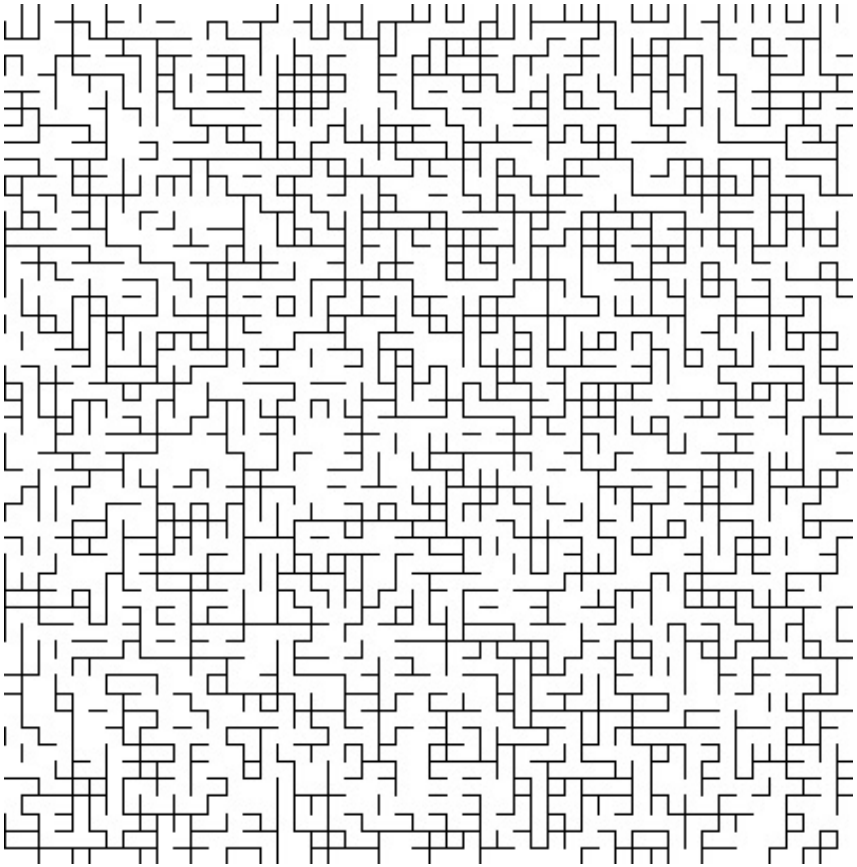


Figure: Planar percolation

Critical percolation in 2D and 3D

Two dimensions, much is known

Russo-Seymour-Welsh; Kesten $p_c = 1/2$; Kesten's bound on influences; Smirnov's conformal invariance theorem; Schramm's SLE; Noise sensitivity; Garban-Pete-Schramm's spectral description for the crossing event in planar percolation;

Three dimensions, little is known

(Famous) Problem 5: Understand 3-D percolation.

- No percolation at p_c?
- Russo-Seymour-Welsh?
- Bounds on influences?
- noise-sensitivity? scaling limit? Description of the Fourier probability distribution?

Theorem (Kalai and Kozma, 2016+): For the crossing event in 3D percolation, either $I \geq N^{\beta}$, $\beta > 0$, or the energy on "well-separated sets S," with $|S| \leq \log n^{3/2}$ is negligible.

Two slides on Projections (Bourgain, Kahn, Kalai, 2014)

For a set $A \subset \Omega_n$ and $T \subset [n]$ let A_T denote the projection of A on the coordinates in T.

Problem 5.1: Let $b, 1/2 < b < 1$ be a real number. For a subset A of the discrete cube of measure $1/2$, what can be said on the maximum value of $\mu(A_T)$, for $T \subset [n]$, and $|T| = bn$.

We now have a good understanding for this problem.

Problem 5.2: Given $b, 1/2 \leq b < 1$ what is the minimum value t so that every set $A \subset \Omega_n$ of measure t has a projection A_T for some subset T, $|T| = bn$ with $\mu(A_T) = 1/2$.

Recent progress both for lower and upper bounds but still a huge gap in our knowledge.

Projections into half the dimension:

Problem 5.3:
Given $\mu(A)$ what can be said about the maximum of $\mu(A_T)$ for $|T| = n/2$.

Remark: The Sauer-Shelah lemma kicks in for projections into subcubes of less than half the dimension.

Correlation and extremal combinatorics

Problem 6: Given two monotone functions f and g what can be said about $cov(f,g)$ the covariance between f and g?
Here, $cov(f,g) = \mathbb{E}(fg) - \mathbb{E}(f)\mathbb{E}(g)$. Harris and Kleitman proved that $cov(f,g) \geq 0$.

Talagrand's correlation formula: If f and g are monotone Boolean functions with n variables then

$$cov(f,g) \geq K \sum_{k=1}^{n} I_k(f) I_k(g) / \log(1/\sum_{k=1}^{n} I_k(f) I_k(g)).$$

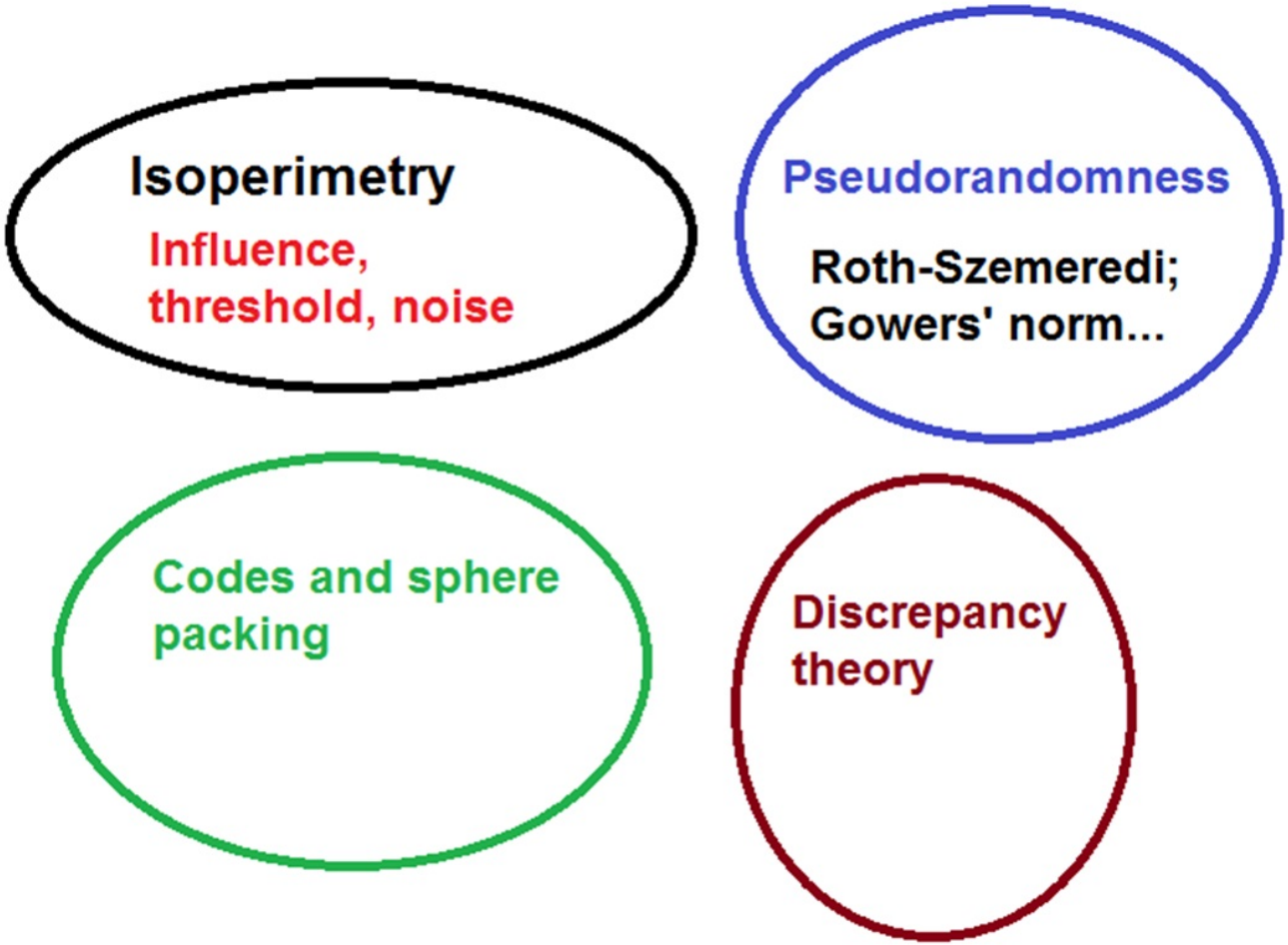
Kahn's correlation conjecture: If f and g are monotone Boolean functions and g is an odd function (i.e., $g(1-x) = 1 - g(x)$) then

$$cov(f,g) \geq \sum_{k=1}^{n} I_k(f) \sum_{S: k \in S} \frac{1}{|S|} \hat{g}^2(S).$$

Computation

The combinatorics, analysis and geometry of Boolean functions, Fourier expansion, and noise, have strong connections with various areas of the theory of computing: algorithms, computational complexity, derandomization, distributed computing, cryptography, error-correction codes, and quantum information and computation.

Fourier Analysis and combinatorics



Generalizations

- Other product spaces
- Other probability distributions
- Other graphs
- Other groups (and representations)
- Other (simple) polytopes
- High dimensional expansions

ims Textbooks
Noise Sensitivity of Boolean Functions and Percolation
Christophe Garban
Jeffrey E. Steif
CAMBRIDGE
Mathematical Surveys and Monographs
Volume 89
Copyrighted Material
The Concentration of Measure Phenomenon
Michel Ledoux
American Mathematical Society
Copyrighted Material
Analysis of Boolean Functions
RYAN O'DONNELL
Thank you very much